Attention in value-based choice as optimal sequential sampling

Frederick Callaway
Department of Psychology
Princeton University
fredcallaway@princeton.edu

Thomas L. Griffiths
Departments of Psychology and Computer Science
Princeton University
tomg@princeton.edu

Abstract

When faced with a decision between several options, people rarely fully consider every alternative. Instead, we direct our attention to the most promising candidates, focusing our limited cognitive resources on evaluating the options that we are most likely to choose. Despite a growing body of empirical work demonstrating the important role of attention in decision making, little is known about how people choose to allocate their attention when making decisions. To address this gap, we cast attention allocation in decision making as a sequential sampling problem, in which an agent iteratively selects from which distribution to sample in order to update her beliefs about the values of the available alternatives. By approximating the optimal solution to this problem, we derive a model in which both the selection and integration of evidence are rational. This model predicts choices and reaction times, as well as sequences of visual fixations.

Applying the model to a ternary-choice dataset, we find that its predictions align well with human data.

Keywords: decision-making; sampling; attention; eye-tracking

Acknowledgements

This research was supported by a grant from Facebook Reality Labs

1 Introduction

Consider a diner at a new restaurant, perusing the menu and trying to decide what she wants to have for dinner. Under the standard model of economic decision making (Kahneman & Tversky, 1979; Rangel, Camerer, & Montague, 2008), she would assign each item a value and then choose the one with maximal value. While her preferences might change day to day (leading to inconsistent choices), she always chooses the item that she likes best at the moment. Unfortunately, this idealization does not capture the experience many of us have. Instead, we undergo a difficult and sometimes lengthy process of weighing the options, initially being drawn to one entrée before identifying a competitor, oscillating between them, failing to notice a desirable choice, and feeling pangs of regret at the sight of our companion's meal.

A great deal of work in psychology and neuroscience has attempted to better capture the process through which people make decisions. One key insight is that decision-making is a sequential process. This insight is formalized in the drift-diffusion model (Mormann, Malmaud, Huth, Koch, & Rangel, 2010; Ratcliff, 1978), where the oscillation between mushroom risotto and pesto gnocchi is captured by the perturbations of a random walk. However such models fail to capture the important role that attention plays in the decision making process (Orquin & Mueller Loose, 2013). This gap is partly addressed by the attentional drift diffusion model (aDDM), in which the drift rate is shifted in favor of the fixated item (Krajbich, Armel, & Rangel, 2010). However, while such models achieve impressive fits to empirically observed relationships between attention and choice, they do not provide a theory of how people choose what to look at.

Here, we attempt to provide a rational model of attention allocation in decision making within the framework of resource-rational analysis (Griffiths, Lieder, & Goodman, 2015). We begin by formalizing the attention allocation problem as a process of sequential sampling and posterior updating. By endogenizing attention within this model (i.e. allowing an agent to select which item to consider at each time point), we define a sequential decision problem *for how to make a single decision*. We formalize this problem as metalevel Markov decision process (Hay, Russell, Tolpin, & Shimony, 2012), and identify an approximately optimal solution using a recently developed reinforcement learning technique (Callaway, Gul, Krueger, Griffiths, & Lieder, 2018). With this approach, we can make precise predictions about both how attention should affect choice as well as how attention should be allocated, all within a single rational model. We identify two key predictions: Attention is preferentially directed to high-value items, and highly-attended items are more frequently chosen. Both of these characteristics have been previously observed in human data, but (to our knowledge) we provide the first rational explanation for the effects. We illustrate the findings with the dataset of Krajbich and Rangel (2011), in which participants made ternary choices while under eye-tracking, finding that the model captures several of the key patterns in the human data.

2 Model

We treat decision making as an iterative process of sampling and inference. The decision maker (or *agent*) is presented with a set of items, each of which has some true unknown value. In order to determine which item to choose, the agent can generate noisy samples of their utilities, each sample providing a small amount of information about the utility of a single item, but also having a small cost. The agent continuously integrates information from these samples, thus developing an increasingly precise and accurate belief about each item's value.

The role of attention in this model is to select which item is sampled at each time step. Importantly, the agent cannot simply allocate attention to the highest value item because she does not know the true values. Rather, she must decide which item to attend to based on her current value estimates. She may choose, for example, to focus her attention on items that she thinks have high value, or perhaps on items whose value is still highly uncertain. However, we do not specify how expected value and uncertainty should be traded off. Instead, we assume that the agent allocates attention in an (approximately) optimal way, where optimality is defined as maximizing decision utility minus computational cost.

Our model is based on theoretical work in artificial intelligence, specifically the field of metareasoning (Russell & Wefald, 1991). It is thus interesting, if not surprising, that the model bears some resemblance to the evidence accumulation models more familiar to psychologists, in particular the drift diffusion model (DDM). As in the DDM, decisions in our model arise from a process in which evidence for and against each item is accumulated over time, being tracked by internal decision variables. However, our model contrasts with the DDM in that both expected values and uncertainty estimates are explicitly represented. This allows the decision variables to evolve according to optimal Bayesian inference.[1] Furthermore, we do not posit a fixed stopping rule such as a threshold, but instead assume that the stopping point is determined on the fly by a dynamic cost-benefit analysis. Finally, unlike any evidence accumulation models of which we are aware, we propose that the information is not gathered either uniformly for all items, or by a stochastic (exogenous) attentional process; instead, information is selected by a near-optimal meta-controller.

[1] Only the binary, hypothesis-testing version of the DDM (e.g. as applied to perceptual decision making) has been shown to implement Bayesian inference. This is in contrast to the multi-alternative, value-based decisions we model here.

2.1 Attention allocation as a Metalevel Markov decision process

A metalevel Markov decision process (meta-MDP) is a formalism developed in the artificial intelligence literature to describe the problem of allocating computational resources to best trade off between decision quality and computational cost (Hay et al., 2012). The key insight lies in viewing computation as a sequential process: an algorithm typically executes many individual operations to accomplish an ultimate goal. Given this observation, it is natural to model computation as a Markov decision process because it is the standard formalism for modeling sequential decision problems.

A meta-MDP is formally identical to a standard MDP; it is defined by a set of states, a set of actions, a transition function, and a reward function. It is distinct from a standard MDP only in its interpretation and the way in which it is derived. In a meta-MDP the states correspond to the agent's beliefs, the actions correspond to computations (or cognitive operations), the transition function describes how computations update the agent's beliefs, and the reward function describes both the cost of computation and also the utility of the item that is ultimately chosen.We now define each of these elements.

Beliefs The agent's beliefs are described by a set of Gaussian distributions, each of which is the agent's posterior distribution for the utility of one item. We denote the posterior utility distribution for item i at time point t as $U_i(t) \sim \mathrm{Normal}(\mu_i(t), \lambda_i(t)^{-1})$. The belief at time step t, $b(t)$, can thus be encoded by two vectors giving the mean, $\mu(t)$, and precision, $\lambda(t)$, of the estimate of each item's value. The state space of the meta-MDP is thus $\mathbb{R}^k \times (0, \infty)^k$.

Computations The actions of the meta-MDP correspond to the computations (or cognitive operations) that the agent can perform. Although decision making likely draws on many types of computation, we consider a highly reduced set that captures attention directed towards each item, $\{c_1, c_2, \dots c_k\}$, where c_i updates the belief about item i, as described in the following paragraph. Additionally, we define a special computation, $\perp$, which indicates that the agent terminates the decision making process and makes the best choice given her current beliefs, choosing an item according to $\arg\max_i \mu_i(t)$.

Transition function The metalevel transition function describes how computations affect beliefs. Each computation draws a sample from an observation distribution for a single item, $\mathrm{Normal}(u_i, \sigma^2)$, with mean equal to the item's true utility. The belief about that item's value is then updated in accordance with Bayesian inference. Let c denote the item considered at time step t. The transition dynamics are then defined by the following equations:

$$\begin{aligned} o(t) \sim \mathrm{Normal}(u_c, \sigma^2) \quad \lambda_c(t+1) &= \lambda_c(t) + \sigma^{-2} & \lambda_i(t+1) &= \lambda_i(t) \text{ for } i \neq c \\ \mu_c(t+1) &= \frac{\sigma^{-2} o(t) + \lambda_c(t)\mu_c(t)}{\lambda_c(t+1)} & \mu_i(t+1) &= \mu_i(t) \text{ for } i \neq c \end{aligned} \tag{1}$$

Reward function The metalevel reward function captures both the costs of computation and also the quality of the decision that is ultimately made. The costs are defined

$$R(b_t, c_t) = -\mathrm{cost}_{\text{sample}} - \mathbf{1}(c_t \neq c_{t-1})\mathrm{cost}_{\text{switch}}\,, \tag{2}$$

where the first term captures the cost of sampling and the second captures an additional switching cost for considering a different item than the one considered on the previous time step. To maintain the Markov property, we simply add the previous computation, c_{t-1}, as an auxiliary state variable. The metalevel reward function captures decision quality by assigning a reward to the termination action, $\perp$. When this action is selected, the agent chooses the best item given her current beliefs, $i^*(t) = \arg\max_i \mu_i(t)$. The reward for terminating is thus $R(b_t, \perp) = u_{i^*(t)}$.

2.2 Optimal attention allocation as an optimal metalevel policy

Having formalized the problem of attention allocation for decision making as a metalevel Markov decision process, the problem of identifying an optimal attentional strategy is reduced to the problem of identifying the optimal policy for an MDP. The policy of a meta-MDP is a function that returns a computation (or distribution over computations) to take in a given belief state, and the metalevel policy is optimal if it maximizes total expected metalevel reward. Because the belief state space is continuous, standard dynamic programming or tabular reinforcement learning techniques cannot be applied. Thus, we employ a recently developed feature-based policy search method that has been shown to find near-optimal policies on meta-MDPs similar to the one presented above (Callaway et al., 2018).

3 Results

To test the model's ability to capture the structure of human attention in value-based choice, we apply it to the dataset of Krajbich & Rangel (2011), in which participants made ternary choices between snack items. Response time was free and visual attention was recorded with eye tracking. For each of the 2966 trials in the dataset, we simulate ten model runs, using the z-scored ratings the participants provided in the first stage of the experiment as the true utilities, u_i, of

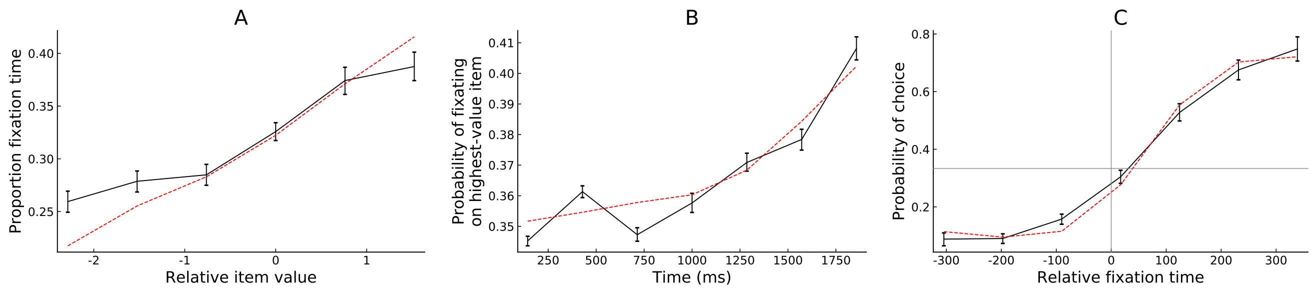

Figure 1: Model predictions. (a) Proportion of total fixation time spent on a given item as a function of its relative value. Relative value is defined as the item's rating minus the mean rating for all options in the given trial. (b) Probability of fixating on the item with maximal value as a function of the time since trial onset. Trials which lasted less than 2 seconds are excluded from this plot so that all plotted points are based on the same set of trials (to avoid confounding time with difficulty). (c) Probability of choosing an item as a function of its relative fixation time. Relative fixation time is defined as the amount of time spent fixating on the given item minus the mean fixation time for all items in the trial. The dashed red lines indicate the model simulation. Bars denote 95% confidence intervals estimated by bootstrapping.

each item. On each of these runs, the model generates a sequence of computations (each corresponding to attending to a single item) as well as a choice. We map the sequence of computations onto visual fixations by assuming that, for each computation, the model fixates on the attended item for t_{sample} milliseconds. t_{sample} is set to match mean predicted reaction time to the empirical mean. The free parameters of the model, σ, $\text{cost}_{\text{sample}}$, and $\text{cost}_{\text{switch}}$ were fit by minimizing mean absolute error on a set of summary statistics.[2] The model makes two key predictions that we discuss in detail: attention is preferentially directed to high-value items, and highly-attended items are more frequently chosen.

3.1 Attention is directed to high-value items

Our model predicts that a rational agent will allocate more attention to items that she believes are valuable. As shown in Figure 1a, humans appear to do the same. To see why this is rational, consider a case in which there are two items with similarly high value and one item with much lower value. The agent can quickly determine that the low value item is not the best one, and thus allocates most of her attention to the two high-value items to determine which one to choose. In the reverse case, in which there is only one high value item, the agent quickly identifies it and has no reason to determine which of the similarly low valued items is superior; attention is roughly equally divided among the three options. Thus, in net, positively valued items receive more attention.

Importantly, the true item values do not directly influence attention; this effect is mediated by the internal value estimates that the agent constructs over the course of a decision. In the initial stages of a decision, the agent's value estimates will are uncertain and noisy, with a weak dependence on the true values. Although she attends most to the items she *believes* are most valuable, she is likely to be mistaken in her belief and thus frequently attends to low-value items. As the decision progresses, the agent refines her belief and the estimated values that are biasing attention become closer to the true values. As a result, we expect that the tendency to attend to high value items will increase over the course of the decision. Indeed, as shown in Figure 1b, this pattern holds in both the human data and the model simulations.

Importantly, this prediction only holds for decisions between three or more items. Indeed, it has been shown that attention should be exactly evenly divided in the two alternative case (Fudenberg, Strack, & Strzalecki, 2018), and our policy optimization method recovers this behavior. Accordingly, we see a relatively weak effect of value on attention in the two-alternative case (Krajbich et al., 2010). This small effect might be due to people applying a heuristic attention-allocation policy which is explicitly biased towards high-value items. Such a strategy could closely approximate optimal attention allocation in the more common multi-alternative case, but result in slightly suboptimal behavior in the two-alternative case. Exploring heuristic strategies that can explain this phenomenon is an important direction for future research.

3.2 More attended items are more likely to be chosen

In our model, the more an agent attends to an item, the more likely she is to ultimately choose it (Figure 1c). This prediction is also made by the aDDM, a direct result of the assumption that the drift rate is biased towards the attended item. However, our model makes the same prediction without any biases in the evidence accumulation process.

[2] Maximum likelihood estimation would be a more principled approach to fitting model parameters. However, the high dimensionality of both the latent belief states and the predicted sequence of fixations makes computing or even approximating the likelihood by standard methods computationally intractable. We are investigating custom inference algorithms to address this challenge.

There are two mechanisms through which the effect can emerge in our model. The first mechanism is through a mismatch between the prior distribution and the true distribution from which items are drawn. Value estimates in our model are based on a combination of a prior and likelihood, where the weight of the likelihood increases with computation time. Thus, for an item with true value above the prior mean, the estimated value tends to increase with computation time, and the converse holds for items with true value below the prior mean (Armel & Rangel, 2008). If the prior is unbiased, then the positive and negative effects of attention on valuation wash out. However, if the prior is biased, i.e. if choice items are systematically better or worse than expected, attention has a corresponding systematic effect on valuation.

The second mechanism by which the apparent attentional bias may emerge depends on the rational allocation of attention. As described in the previous section, the rational model predicts that (in the case of three or more items), the agent is more likely to attend to items that she believes are valuable. At the same time, she is also more likely to choose items that she believes are valuable. Thus, a high estimated value makes the agent more likely to both attend to and choose an item. As a result, attention and choice are correlated by a common-cause structure. This contrasts to the mechanism described above in which attention has a causal effect on choice, as mediated by estimated value. Importantly, both choice and attention allocation depend on estimated value, not the true item values; thus, the correlation is not broken by conditioning on true value.

4 Discussion

We presented a rational model of attention allocation in value-based decision making that formalizes attention allocation as a sequential sampling problem. Here, we focused on two key predictions of the model: attention is directed to high-value items and more attended items are more likely to be chosen. However, because the model predicts full sequences of fixations, we can look for additional model predictions in the same way that we analyze rich and complex human data, through exploratory data analysis. Preliminary such analyses have revealed, for example, that the model predicts longer reaction times for trials in which all items are relatively bad, perhaps reflecting a form of loss attention (Yechiam & Hochman, 2013). Further investigating the model's predictions is a key direction for future work.

In the results presented here, we normalized the ratings such that the model's prior is unbiased, ruling out the biased-prior explanation for the correlation between attention and choice. Thus, Figure 1a seems to suggest that this correlation (at least as it appears in this dataset) can be explained entirely by a common-cause structure in which high estimated values lead to both more attention and greater choice probability. However, experimental manipulations of attention in a perceptual decision-making task suggest that there is in fact a causal effect of attention on choice (Tavares, Perona, & Rangel, 2017). In further work, we plan to apply our model to this and other datasets to further elucidate the role attention plays in decision making.

References

Armel, K. C., & Rangel, A. (2008). Neuroeconomic models of economic decision making: The impact of computation time and experience on decision values. *American Economic Review*, *98*(2), 163–168.

Callaway, F., Gul, S., Krueger, P., Griffiths, T. L., & Lieder, F. (2018). Learning to select computations. In *Uncertainty in artificial intelligence: Proceedings of the thirty-fourth conference.*

Fudenberg, D., Strack, P., & Strzalecki, T. (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review*, *108*(12), 3651–3684.

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, *7*(2), 217–229.

Hay, N., Russell, S., Tolpin, D., & Shimony, S. (2012, aug). Selecting Computations: Theory and Applications. *Uncertainty in Artificial Intelligence: Proceedings of the Twenty-Eighth Conference*.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298.

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, *108*(33), 13852–13857.

Mormann, M., Malmaud, J., Huth, A., Koch, C., & Rangel, A. (2010, 10). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgment and Decision Making*, *5*(6), 437–449.

Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica*, *144*(1), 190–206.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545–556.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59–108.

Russell, S., & Wefald, E. (1991). Principles of metareasoning. *Artificial Intelligence*, *49*(1-3), 361–395.

Tavares, G., Perona, P., & Rangel, A. (2017, Aug). The attentional drift diffusion model of simple perceptual decision-making. *Frontiers in Neuroscience*, *11*.

Yechiam, E., & Hochman, G. (2013, Mar). Loss-aversion or loss-attention: The impact of losses on cognitive performance. *Cognitive Psychology*, *66*(2), 212–231.